REMARKS

The Office Action and the cited and applied references have been carefully studied. Claims 1, 2 and 4-9 are allowed. Claims 1, 2, 4-9, and 18-52 presently appear in this application and define patentable subject matter warranting their allowance. Reconsideration and allowance are hereby respectfully solicited.

Attached hereto are copies of Refs. AN (XP002024314) and BG (Japan Abstract 05279376), which appear to be the same, requested by the examiner in Paper No. 12 to forward prosecution because the references were not found in USPTO files.

Claims 18, 20, 21-57 have been rejected under 35 U.S.C. §112, first paragraph, because the examiner states that the specification, while being enabling for a composition comprising SEQ ID NO:6, or for derivatives thereof varying from SEQ ID NO:6 by one amino acid residue, does not reasonably provide enablement for a composition comprising any homologue of the sequence. This rejection is respectfully traversed.

As the examiner stated, the present specification discloses two interferon-gamma inducing polypeptides. One of them has the amino acid sequence of SEQ ID NO:6, wherein amino acid residue 73 is Ile, and the other has the amino acid sequence of SEQ ID NO:6 wherein the amino acid residue 73 is Thr. The specification from page 9, line 13 to page 10, line 15, further discloses a polypeptide as defined in claims 18 and 20, which has

a "homologous sequence" of SEQ ID NO:6. Applicants therefore believe that the subject matter as defined in claims 18 and 20 are literally supported by the specification.

Furthermore, even though the specification discloses only two kinds of "homologous sequences", one of skill in the art could easily obtain other "homologous sequences" of SEQ ID NO:6 based on the information about the consensus sequence between SEQ ID NO:6 and SEQ ID NO:4 and being guided by the molecular weight, isoelectric point and biological activity as defined in claim 18.

The examiner states in the paragraph bridging pages 5 and 6 of the Office Action that while a skilled person may feel that the sequence region comprising residues 97-113 is likely to be necessary for protein function, an IGIF homologue and an isoform thereof described in WO 98/10072 lack the region, and thus, what one of ordinary skill in the art may have recognized as necessary for protein function from the presently disclosed sequences is unpredictably disclosed as unnecessary. With due respect to the examiner, the examiner's position here is incorrect. It is an isoform of rat IGIF, i.e., IL-18α, disclosed at pages 7-8 of WO '072 that lacks the consensus sequence (the sequence region comprising residues 97-113). It should be noted however that the rat IL-18 (as opposed to rat IL-18α) disclosed at page 6 of WO '072 having an amino acid sequence of SEQ ID NO:2 does indeed comprise the consensus sequence (the sequence regions

comprising residues 97-113). Accordingly, it is believed that WO '072 cannot be used as evidence to deny that the sequence region comprising residues 97-113 as a consensus sequence between SEQ ID NO:6 and SEQ ID NOA is necessary for protein function.

Moreover, the fact that an isoform of rat IGIF disclosed in WO '072 lacks the sequence region comprising residues 97-113 indicates that other isoforms can be exist in which even a part of the region may be replaced with other amino acid residues.

Attached hereto is a Schematic Diagram of an amino acid alignment between the four amino acid sequences, i.e., SEQ ID NO:6 (human IL-18), SEQ ID NO:4 (mouse IL-18), SEQ ID NO:2 of WO '072 (rat IL-18) and SEQ ID NO:4 of WO '072 (rat IL-18α). Please note that amino acid residues 1-36 of SEQ ID NO:2 of WO '072 (rat IL-18) and SEQ ID NO: 4 of WO '072 (rat IL- 18α) have been excluded from this comparison because they are leader peptides and not part of the mature protein. From the attached Schematic Diagram, it is understood that the four amino acid sequences (i.e. human IL-18, mouse IL-18, rat IL-18 and rat IL-18α (an isoform of rat IL-18)) have several consensus sequences. With regard to consensus sequences, it should be noted that a technique for obtaining a polypeptide having a certain function by binding plural consensus sequences had been known to those of skill in the art at the time the present

application was filed. Derivatives of interferon-α obtained in the beginning of the 1980's are some examples (See for example U.S. Patent Nos. 4,695,623; 5,541,293; 5,661,009).

In short, even if the specification does not teach which amino acid residues of SEQ ID NO:6 can be replaced with other amino acid residues to obtain homologues of human IL-18, one of skill in the art would be able to readily obtain various homologues of SEQ ID NO:6 (i.e., human IL-18) based on the teaching in the specification from page 9, line 13 to page 10, line 15, along with information about consensus sequences between SEQ ID NO:6 (human IL-18) and SEQ ID NO:4 (mouse IL-18) and the state of the art at the time the application was filed. The polypeptide as defined in claim 18 can be easily screened from the various homologues of SEQ ID NO:6 obtained, in the light of physicochemical properties (2) to (4) as recited in claim 18. Only routine experimentation is required to obtain the homologues.

Reconsideration and withdrawal of the rejection are therefore respectfully requested.

Claims 18, 20, 21-52, and 53-57 have been rejected under 35 U.S.C. §112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had

possession of the claimed invention. This rejection is respectfully traversed.

Applicants point out that additional homologues are disclosed in the specification. For instance, Example 7-3, on page 56, discloses a homologue in which two amino acid residues are added to the N-terminus of the amino acid sequence of SEQ ID NO:6, and on page 89, lines 16-19, a homologue with one amino acid residue added to the N-terminus of the amino acid sequence of SEQ ID NO:6 is disclosed.

In addition, the applicants believe that the state of the art at the time the present application was filed should be taken into account when considering if the subject matter was described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors had possession of the claimed invention. As to the state of the art, the applicants would like to point to the publications cited at page 34, first paragraph, of the present specification and U.S. Patent Nos. 5,789,199 and 5,878,373, relevant pages of which are attached hereto. A copy of another publication, "MOLECULAR BIOLOGY OF THE GENE", by James D. Watson et al., The Benjamin/Cummings Publishing Company, Inc., 1987, pp. 222-231, 444-446 is also attached hereto to show the state of the art at the time the present application was filed. From these publications, it would be easily understood that it was routine

work for one of skill in the art at the time the present invention was made to prepare homologues retaining the same biological activity as the original polypeptide if the amino acid sequence of the polypeptide is given.

The present specification does indeed disclose some examples of homologues as well as the amino acid sequence of the original polypeptide. Applicants therefore believe that the subject matter of the claimed invention was described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors had possession of the claimed invention.

Reconsideration and withdrawal of the rejection are therefore respectfully requested.

Claims 18, 20, 21, 24, 27 and 28 have been rejected under 35 U.S.C. §102(e) as being anticipated by U.S. Patent No. 5,912,324, issued to Okamura et al. This rejection is respectfully traversed.

The mouse IL-18 disclosed in U.S. Patent '324 as can be seen from the comparison/amino acid sequence alignment (Schematic Diagram) attached hereto as SEQ ID NO:4 (sequence labeled B) in the present application (and as SEQ ID NO:2 in '324), has several internal deletions or insertions (not merely substitutions) relative to the human IL-18 amino acid sequence of SEQ ID NO:6 (sequence labeled A). However, as recited in claim 18, additions

and deletions are only made from the N-terminus and/or C-terminus. Therefore, the mouse IL-18 of U.S. Patent '324 is excluded from being a homologue of SEQ ID NO:6 as defined in claim 18 and U.S. Patent '324 cannot anticipate the presently claimed invention.

Reconsideration and withdrawal of the rejection are therefore respectfully requested.

Claims 18 and 20 have been rejected under 35 U.S.C. §102(e) as being anticipated by Joh et al., WO 98/10072. This rejection is respectfully traversed.

Similar to what is discussed above in the §102(e) rejection over U.S. Patent '324, Joh discloses rat IL-18 and rat IL-18α, which are presented as sequences C and D, respectively, in the comparison/amino acid sequence alignment (Schematic Diagram) attached hereto and which clearly have internal additions or deletions relative to the human IL-18 amino acid sequence of SEQ ID NO:6 (sequence labeled A). As such internal additions and deletions are excluded from the homologues rejected in claim 18, Joh et al. also cannot anticipate the presently claimed invention.

Reconsideration and withdrawal of the rejection are therefore respectfully requested.

Claims 21-23, 25, 26, and 28-52 have been rejected under 35 U.S.C. §103(a) as being unpatentable over U.S. Patent

No. 5,912,324, issued to Okamura et al. This rejection is respectfully traversed.

The present application, which was filed after November 29, 1999, is commonly owned with U.S. Patent No. '324. Pursuant to 35 U.S.C. §103(c), a U.S. Patent that is commonly owned with an application filed after November 29, 1999, is not available as prior art under 35 U.S.C. §102(e)/103(a) against that application. Therefore, U.S. Patent '324 is not available as prior art under §103(a) and cannot make obvious the presently claimed invention.

Reconsideration and withdrawal of the rejection are therefore respectfully requested.

Claims 18, 20, 21, and 26 have been rejected under 35 U.S.C. §103(a) as being unpatentable over Zhou et al., J. Immunol. 155:785-795, in view of Okamura et al., Nature 378:88-91. This rejection is respectfully traversed.

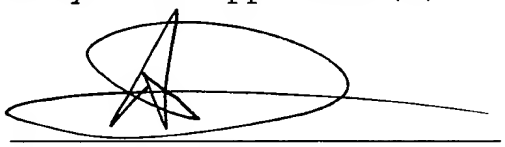
The homologue taught by Okamura is the same mouse IL-18 as disclosed in U.S. Patent '324 discussed above in the §102(e) rejection over '324. Accordingly, such a mouse homologue cannot lead one of ordinary skill in the art to the presently claimed homologues which do not have any internal additions or deletions relative to the human IL-18 of SEQ ID NO:6.

Reconsideration and withdrawal of the rejection are therefore respectfully requested.

In view of the above, the claims comply with 35 U.S.C. §112 and define patentable subject matter warranting their allowance. Favorable consideration and early allowance are earnestly urged.

Respectfully submitted,

BROWDY AND NEIMARK, P.L.L.C.
Attorneys for Applicant(s)

By 
Allen C. Yun
Registration No. 37,971

ACY:pp
Telephone No.: (202) 628-5197
Facsimile No.: (202) 737-3528
G:\BN\S\SUMA\USHIO2\pto\AMD OA 5-19-03.doc

AN

XP 002024314

1/1 - (C) WPI / DERWENT
AN - 93-374598 ç47!
AP - JP910222423 910808
PR - JP910222423 910808
TI - IFN-gamma induction active substance having in=vivo antitumour activity - obtd. by sepg. antigen substance from Streptomyces haemolyticus prepn., purifying using affinity chromatography etc.
it - GAMMA INDUCTION ACTIVE SUBSTANCE IN-VIVO ANTITUMOUR ACTIVE OBTAIN SEPARATE ANTIGEN SUBSTANCE STREPTOMYCES HAEMOLYTICUS PREPARATION PURIFICATION AFFINITY CHROMATOGRAPHY
PA - (SATO/) SATO M
PN - JP5279376 A 931026 DW9347 C07G17/00 006pp
ORD - 1993-10-26
IC - A61K37/66 ; C07G17/00 ; C12N5/20 ; C12N15/06 ; C12P21/08 ; (C12P21/08 C12R1:91)
FS - CPI
DC - B04 D16
AB - J05279376 INF-gamma induction active substance (I) has the following properties: (1) mol.wt. of 70 kilodalton; and (2) reactivity with monoclonal antibody TS-2. The inducing activity of IFN-gamma is not inactivated by (1) heat treatment of 56 deg.C for 30 mins., (2) treatment by Pronase (0.2 mg/ml) at 37 deg.C for 1 hr., (3) treatment by Neuraminaidase (0.5 IU/ml) at 37 deg.C for 2 hrs., or (4) treatment by 0.15M NaCl-ammonia (pH 11); it is inactivated by (5) treatment by 50 mM periodic acid at 4 deg.C for 2 hrs., and (6) treatment by 0.2M glycine-HCl (pH 2.5). In the prepn. of (I) from a glycolipid specimen extracted from Streptococcus haemolyticus prepn. (OK-432) using butanol by Morrison's method, an antigen substance TS-2 recognised by a monoclonal antibody against OK-432, is sepd. and purified by using an affinity chromatography contg. TS-2 as the ligand. The purified specimen is fractionated to 25 fractions in a Sephacryl S-300 column and the substance corresp. to fraction 8 is collected.
- USE/ADVANTAGE - The new substance has antitumour activity in vivo.
- In an example, a soln. of 50 kg of OK-432 in 5 ml physiological saline soln. was mixed with 5 ml 1-butanol. The mixt. was centrifuged. 20 microg/ml of Pronase was added to the aq. layer and reacted at 37 deg.C for 24 hrs. The reaction mixt. was centrifuged and the supernatant was dialysed against PBS to give a glycolipid specimen (OK-PS). Monoclonal antibody TS-2 was prepd. from Balb/c mouse immuned by OK-432. Antigen substance recognised by TS-2 was sepd. and purified to give OK-PSA. OK-PSA was fractionated to 25 fractions. Fraction 8 extends the living period of mouse infected by HSG cell. (Dwg.0/0)

BNSDOCID: <XP 2024314A>

BG

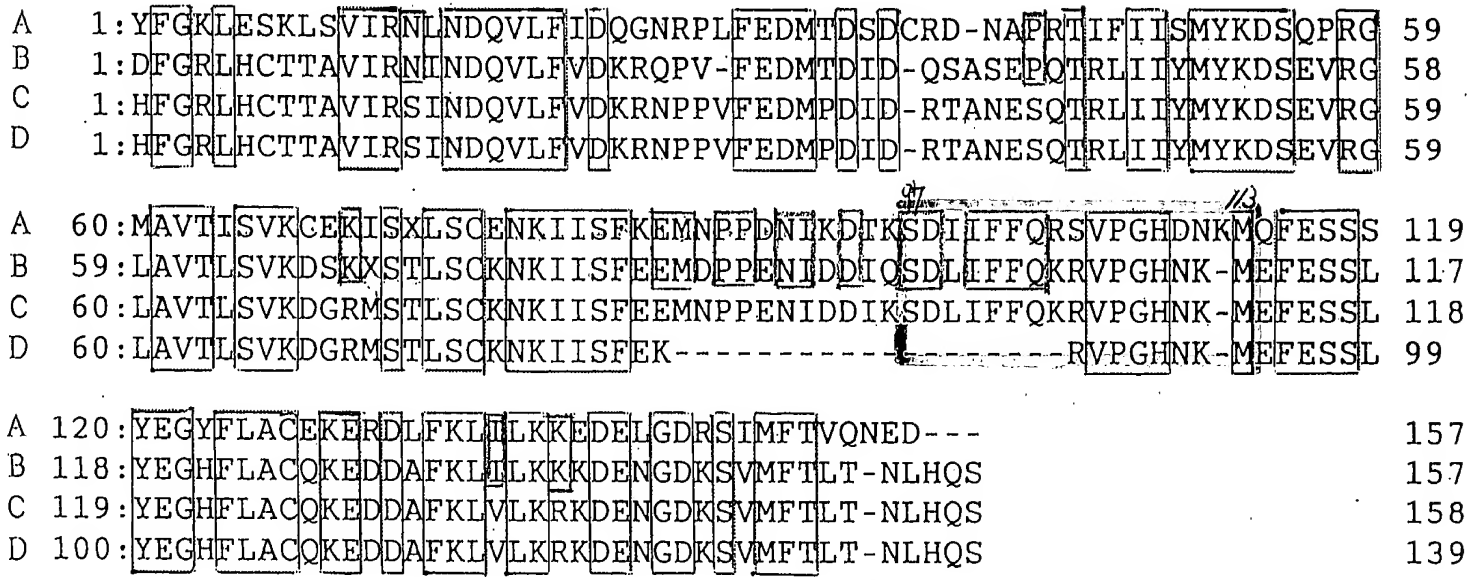
XP. 2024314

1/1 - (C) WPI / DERWENT

AN - 93-374598 [47]
AP - JP910222423 910808
PR - JP910222423 910808
TI - IFN-gamma induction active substance having in=vivo antitumour activity - obtd. by sepg. antigen substance from Streptomyces haemolyticus prepn., purifying using affinity chromatography etc.
it - GAMMA INDUCTION ACTIVE SUBSTANCE IN-VIVO ANTITUMOUR ACTIVE OBTAIN SEPARATE ANTIGEN SUBSTANCE STREPTOMYCES HAEMOLYTICUS PREPARATION PURIFICATION AFFINITY CHROMATOGRAPHY
PA - (SATO/) SATO M
PN - JP5279376 A 931026 DW9347 C07G17/00 006pp
IC - A61K37/66 ; C07G17/00 ; C12N5/20 ; C12N15/06 ; C12P21/08 ; (C12P21/08 C12R1:91)
AB - J05279376 INF-gamma induction active substance (I) has the following properties: (1) mol.wt. of 70 kilodalton; and (2) reactivity with monoclonal antibody TS-2. The inducing activity of IFN-gamma is not inactivated by (1) heat treatment of 56 deg.C for 30 mins., (2) treatment by Pronase (0.2 mg/ml) at 37 deg.C for 1 hr., (3) treatment by Neuraminaidase (0.5 IU/ml) at 37 deg.C for 2 hrs., or (4) treatment by 0.15M NaCl-ammonia (pH 11); it is inactivated by (5) treatment by 50 mM periodic acid at 4 deg.C for 2 hrs., and (6) treatment by 0.2M glycine-HCl (pH 2.5). In the prepn. of (I) from a glycolipid specimen extracted from Streptococcus haemolyticus prepn. (OK-432) using butanol by Morrison's method, an antigen substance TS-2 recognised by a monoclonal antibody against OK-432, is sepd. and purified by using an affinity chromatography contg. TS-2 as the ligand. The purified specimen is fractionated to 25 fractions in a Sephacryl S-300 column and the substance corresp. to fraction 8 is collected.
- USE/ADVANTAGE - The new substance has antitumour activity in vivo.
- In an example, a soln. of 50 kg of OK-432 in 5 ml physiological saline soln. was mixed with 5 ml 1-butanol. The mixt. was centrifuged. 20 microg/ml of Pronase was added to the aq. layer and reacted at 37 deg.C for 24 hrs. The reaction mixt. was centrifuged and the supernatant was dialysed against PBS to give a glycolipid specimen (OK-PS). Monoclonal antibody TS-2 was prepd. from Balb/c mouse immuned by OK-432. Antigen substance recognised by TS-2 was sepd. and purified to give OK-PSA. OK-PSA was fractionated to 25 fractions. Fraction 8 extends the living period of mouse infected by HSG cell. (Dwg.0/0)

Comparison of amino acid sequence homologies of

human IL-18, mouse IL-18, rat IL-18, and rat IL-18α



A: Human IL-18 (SEQ ID NO:6 in the present specification)

B: Mouse IL-18 (SEQ ID NO:4 in the present specification)

C: Rat IL-18 (SEQ ID NO:2 excluding 1-36 amino acid residues in WO 98/10072)

D: Rat IL-18α (SEQ ID NO:4 excluding 1-36 amino acid residues in WO 98/10072)

VOLUME I GENERAL PRINCIPLES

MOLECULAR BIOLOGY OF THE GENE

FOURTH EDITION

James D. Watson — COLD SPRING HARBOR LABORATORY
Nancy H. Hopkins — MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Jeffrey W. Roberts — CORNELL UNIVERSITY
Joan Argetsinger Steitz — YALE UNIVERSITY
Alan M. Weiner — YALE UNIVERSITY

The Benjamin/Cummings Publishing Company, Inc.
Menlo Park, California · Reading, Massachusetts · Don Mills, Ontario
Wokingham, U.K. · Amsterdam · Sydney · Singapore
Tokyo · Madrid · Bogota · Santiago · San Juan

Cover art is a computer-generated image of DNA interacting with the Cro repressor protein of bacteriophage λ. The image was prepared by the Graphic Systems Research Group at the IBM U.K. Scientific Centre.

Editor: Jane Reece Gillen
Production Supervisor: Karen K. Gulliver
Editorial Production Supervisor: Betsy Dilernia
Cover and Interior Designer: Gary A. Head
Contributing Designers: Detta Penna, Michael Rogondino
Copy Editor: Janet Greenblatt
Art Coordinator: Pat Waldo
Art Director and Principal Artist: Georg Klatt
Contributing Artists: Joan Carol, Cyndie Clark-Huegel, Barbara Cousins, Cecile Duray-Bito, Jack Tandy, Carol Verbeek, John and Judy Waller

Copyright © 1965, 1970, 1976, 1987 by The Benjamin/Cummings Publishing Company, Inc.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America. Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data
Molecular biology of the gene.
Rev. ed. of: Molecular biology of the gene / James D. Watson. 3rd ed. c1976.
Bibliography
Includes index.
Contents: v. 1. General principles.
1. Molecular biology. 2. Molecular genetics.
I. Watson, James D., 1928– . [DNLM: 1. Cytogenetics. 2. Molecular Biology. QH 506 M7191]
QH506.M6627 1987 574.87'328 86-24500
ISBN 0-8053-9612-8

ABCDEFGHIJ-MU-89876

The Benjamin/Cummings Publishing Company, Inc.
2727 Sand Hill Road
Menlo Park, California 94025

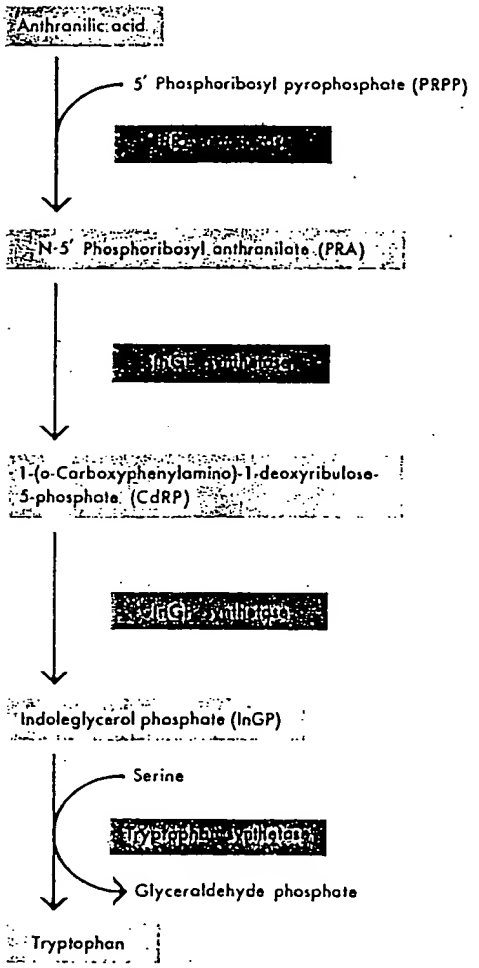


Figure 8-10
Last steps in the pathway of tryptophan biosynthesis.

tial biological activity (they are said to be leaky) and thus are not as easy to work with as those that produce either no product or a drastically rearranged product.

For a long time, the identity of the two proteins coded by the *rIIA* and *rIIB* genes of phage T4 remained unknown. Now they are known to be membrane proteins of molecular weights 86,000 (*rIIA*) and 30,000 (*rIIB*), both present in minor amounts. Unfortunately, still unclear are their exact metabolic functions within T4-infected cells. Thus, for many years there has been a strong tendency to restrict intensive genetic analysis to those genes whose protein products are easy to isolate and whose metabolic or structural roles are well established.

Recessive Genes Frequently Do Not Produce Functional Products

Most mutant genes are recessive with respect to wild-type genes. This fact, puzzling to early geneticists, is now partially understood in terms of the gene-protein relationship. The recessive phenotype often results from the failure of mutant genes to produce any functional protein (enzyme). In heterozygotes, however, there is often present one "good" gene and, correspondingly, a number of "good" gene products. Because the wild-type gene is present in only one copy in heterozygotes, it is possible that there are always fewer good copies of the relevant protein in heterozygotes than in individuals with two wild-type genes. If this were the case, we might guess that the heterozygous phenotype would tend to be intermediate between the two homozygous phenotypes. Usually, however, this does not happen for one of two reasons. Either there are still enough good enzyme molecules to catalyze the metabolic reaction of concern, even though the total number of molecules is reduced, or the recessive gene is not noticeable because control mechanisms cause the single wild-type gene in a heterozygote to produce more gene product than does each wild-type gene in a homozygote. In Chapter 16, we shall discuss how the rates at which bacterial genes act are controlled.

Colinearity of the Gene and Its Polypeptide Product[7]

The best-understood example of the relationship between the order of the mutable sites in a gene and the order of their corresponding amino acid replacements involves the *E. coli* enzyme tryptophan synthetase, one of the several enzymes involved in tryptophan synthesis (Figure 8-10). This enzyme consists of two easily separated polypeptide chains, A and B, neither of which is enzymatically active by itself. A large number of mutants unable to synthesize tryptophan lack a functional A chain in their tryptophan synthetase molecules. When these mutants were genetically analyzed, it was found that changes at a large number of different mutable sites could give rise to inactive A chains. Accurate mapping of these mutants revealed that they all could be unambiguously located on the linear genetic map shown in Figure 8-11. It was possible to isolate the inactive A chains from many of these mutants and to begin to compare their amino acid sequences with the sequence of the wild-type A chain, which contains 267

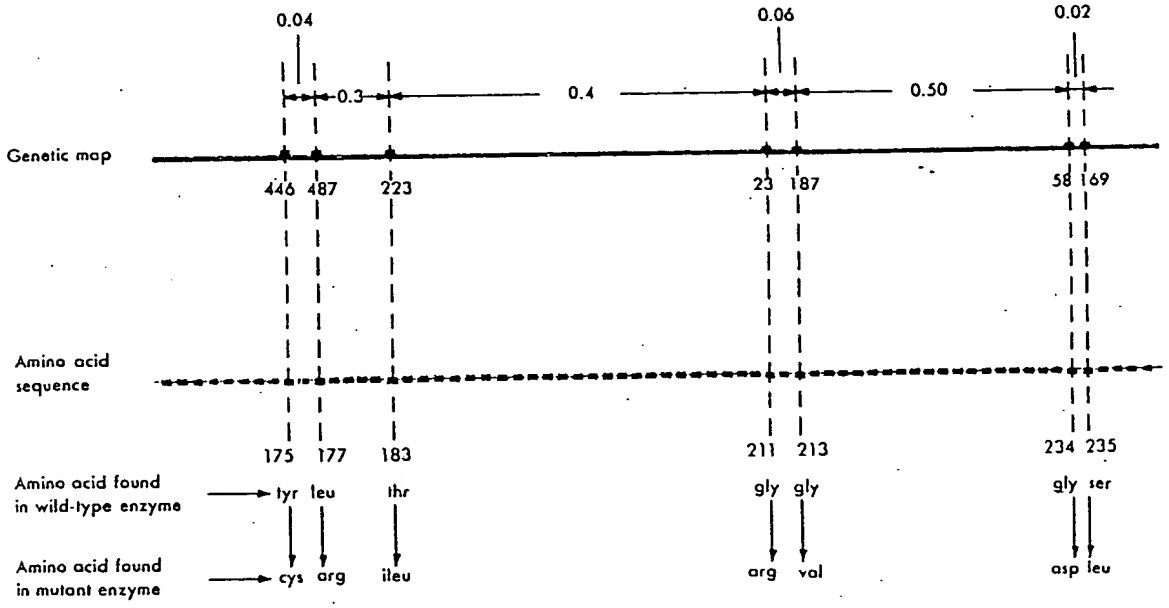


Figure 8-11
Colinearity of the gene and its protein product. Here is the genetic map for one-fourth of the gene coding for the amino acid sequences in the *E. coli* protein tryptophan synthetase A. The designation 0.04, for example, refers to map distances (frequencies of recombination) between tryptophan synthetase mutations A446 and A487. The numbers in the amino acid sequence refer to their position in the 267 residues of the A protein. Following convention, the amino terminal end of the segment is on the left.

amino acids. This sequence allows us to see how the location of a mutation within a gene is correlated with the location of the replaced amino acid in its polypeptide chain product. Since both genes and polypeptide chains are linear, the simplest hypothesis is that amino acid replacements are in the same relative order as the mutationally altered sites in the corresponding mutant genes. This was most pleasingly demonstrated in 1964. The location of each specific amino acid replacement is exactly correlated with its location along the genetic map, a property called colinearity. Thus, successive amino acids in a polypeptide chain are controlled, or coded, by successive regions of a gene.

Mutable Sites Are the Base Pairs Along the Double Helix

In all bacterial genes extensively mapped, the large number of linearly arranged mutable sites that have been found in each gene, and between which genetic recombination (crossing over) is possible, leaves us no choice but to conclude that these sites are the specific base pairs along the DNA of the respective gene (Figure 8-12). A given mutable site can thus exist in any of four different states, AT, TA, GC, or CG. Many mutations are therefore likely to represent simple switches from one state to another. The genetic data that reveal deletions and insertions of genetic material must now be thought of in terms of the addition or deletion of discrete blocks of one to very many base pairs. The three classes of mutations resulting from changes in the sequence of nucleotide bases are illustrated in Figure 8-13.

By carefully studying the fine details of genetic maps, we should be able to obtain important information about the corresponding DNA. However, not every change in base sequence leads to easily observed changes in the corresponding protein. In the genetic code, many amino acids are specified by more than one codon (set of three adja-

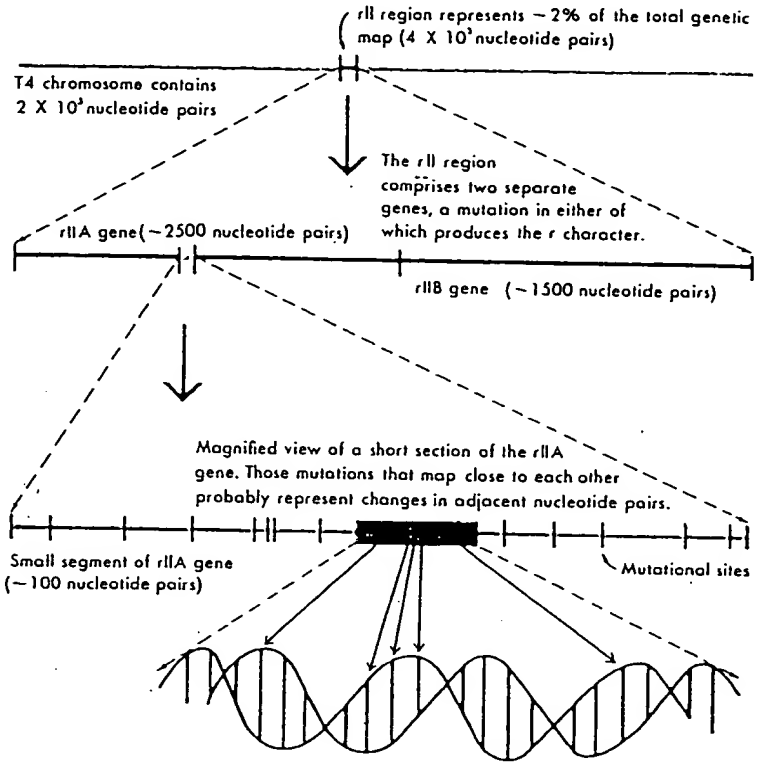


Figure 8-12
The relationship of mutations in the *rII* region of the phage T4 chromosome to the structure of DNA.

cent bases), which means that in many cases, base-pair substitutions will not lead to any amino acid replacements. Moreover, as we document later, many of the amino acids in proteins are not essential, and when they are replaced by somewhat similar amino acids, the proteins often retain full activity. The number of observed mutable sites therefore seriously underrepresents the number of base pairs within the corresponding gene.

There Are Four Alternative Structures for Each Mutable Site[8, 9]

As anticipated, enzymatically inactive tryptophan synthetase molecules resulting from independent mutations at the same mutable site (as shown by failure to give wild-type recombinants) do not always contain the same amino acid replacement. For example, changes in a single mutable site that specifies the amino acid at position 213 results in the replacement of glycine by either glutamic acid or valine. Inspection of the genetic code (see Chapter 15) indicates that in the wild-type strain, this glycine must be specified by either GGA or GGG codons and that the mutable site under study specifies the G in the middle position of this codon. When this G is replaced by U, valine (GUA or GUG) becomes inserted into the glycine site while its replacement by A generates the glutamic acid (GAA or GAG) substitution. Further study of this particular mutable site might eventually turn up the anticipated third replacement in which a G to C switch leads to the appearance of alanine (GCA or GCG).

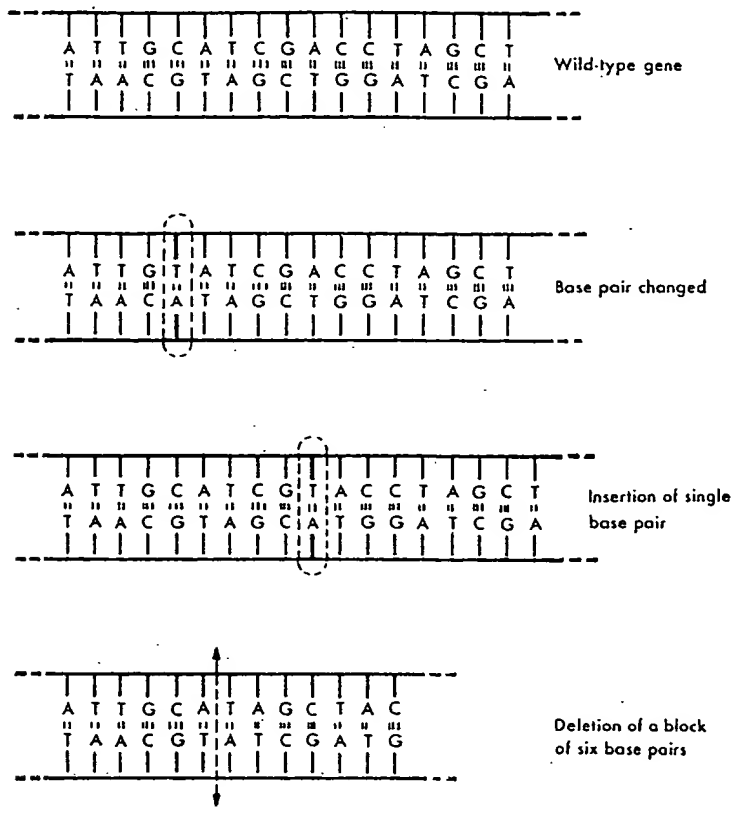


Figure 8-13
Three classes of mutations result from introducing defects in the sequence of bases (A, T, G, C) attached to the backbone of the DNA molecule. In one class, a base pair is simply changed from one into another (i.e., GC to AT). In the second class, a base pair is inserted (or deleted). In the third class, a block of base pairs is deleted (or inserted).

Single Amino Acids Are Specified by Several Adjacent Nucleotide Bases

We expected to find that given amino acids within a particular protein are specified by adjacent mutable sites. This point was first demonstrated in the tryptophan synthetase A gene, where the relevant evidence came from study of the tryptophan synthetase fragment illustrated in Figure 8-14. Treatment of the wild-type strain with a mutagen had given rise to mutant A23, in which arginine replaces glycine (this time at position 212), and mutant A46, in which glutamic acid replaces glycine at the same position. The difference between A23 and A46 does not represent changes to alternative forms of the same mutable site, since a genetic cross between A23 and A46 yields a number of wild-type recombinants (glycine in position 212). If these changes were at the same mutable site, no wild-type recombinants would be produced. Moreover, the very low observed frequency of the wild-type recombinants is compatible with the prediction from the genetic code that these mutable sites are adjacent to each other.

Additional genetic evidence that confirms the separate locations of the A23 and A46 mutable sites comes from observing how A23 and A46 themselves mutate upon treatment with mutagens. After exposure to a mutagen, both strains give rise to new strains, some of which contain active tryptophan synthetase A chains with glycine in position 212. These reverse mutations most likely involve changing the altered mutable sites back to the original wild-type configuration. However, strains containing active tryptophan synthetase also arise

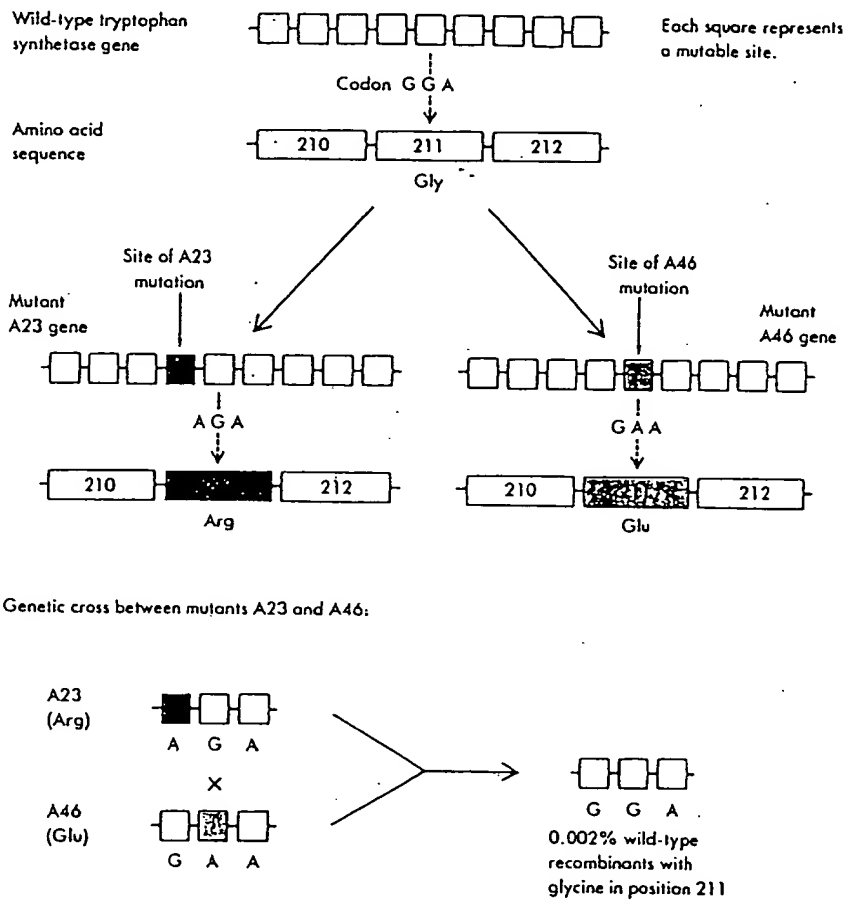


Figure 8-14
Demonstration that a single amino acid is specified by more than one mutable site. We now know that the mutable sites are DNA bases and the codons are actually bases complementary to these in mRNA. (After Emanual J. Murgola.)

in which the amino acid in position 212 is replaced by another amino acid. Most significantly, the type of replacement differs for strains A23 and A46. Besides back-mutating to glycine, strain A23 mutates to threonine and serine, whereas A46 mutates to alanine and valine in addition to glycine. The failure of A23 ever to give rise to alanine or valine and the failure of A46 ever to mutate to threonine or serine is very difficult to explain if their differences from wild type are based on alternative configurations of the same mutable site. But these mutational patterns make perfect sense if glycine at the 212 position is coded by GGA with the A23 mutation to arginine representing a G to A change at the first position of the codon to give rise to AGA and the A46 mutation to glutamic acid occurring at the middle (second) position to give rise to GAA. Their divergent subsequent mutations to serine and threonine and to alanine and valine, respectively, can also be understood by inspecting the genetic code (Figure 8-15).

Single Amino Acid Substitutions Usually Do Not Alter Enzyme Activity

The ability of a polypeptide chain to be enzymatically active does not require an exactly specified amino acid sequence. This is shown by examination of the new mutant strains obtained by treating strains A23 and A46 with mutagens. The possession of either glycine or serine in position 212 yields a fully active enzyme, whereas threonine in

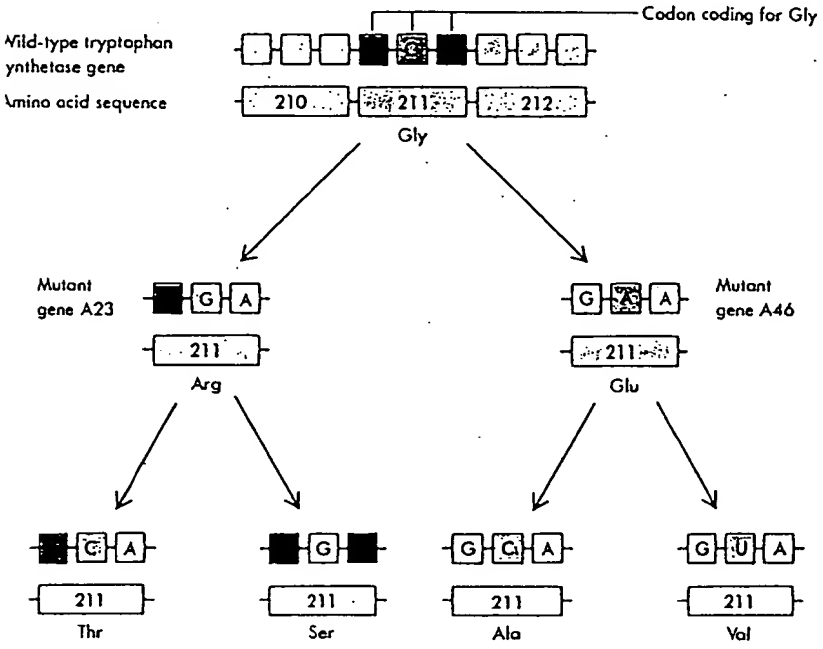


Figure 8-15
Formation of mutants A23 and A46 and their subsequent mutations. Notice that Thr and Ser cannot result from a single base change to the codon for Glu; likewise, Ala and Val cannot result from only one base change to the codon for Arg. Therefore, the A23 and A46 mutants must occur from mutations at two different mutable sites, as shown in Figure 8-14.

the same position yields an enzyme with reduced activity, demonstrating that the activity of an enzyme does not demand a perfectly unique amino acid sequence (Figure 8-16). In fact, evidence now indicates that amino acid replacements in many parts of a polypeptide chain can occur without seriously modifying catalytic activity. However, one sequence may often be best suited to a cell's particular needs, and it is this sequence that is encoded by the wild-type allele. Even though other sequences are almost as good, they will tend to be selected against in evolution.

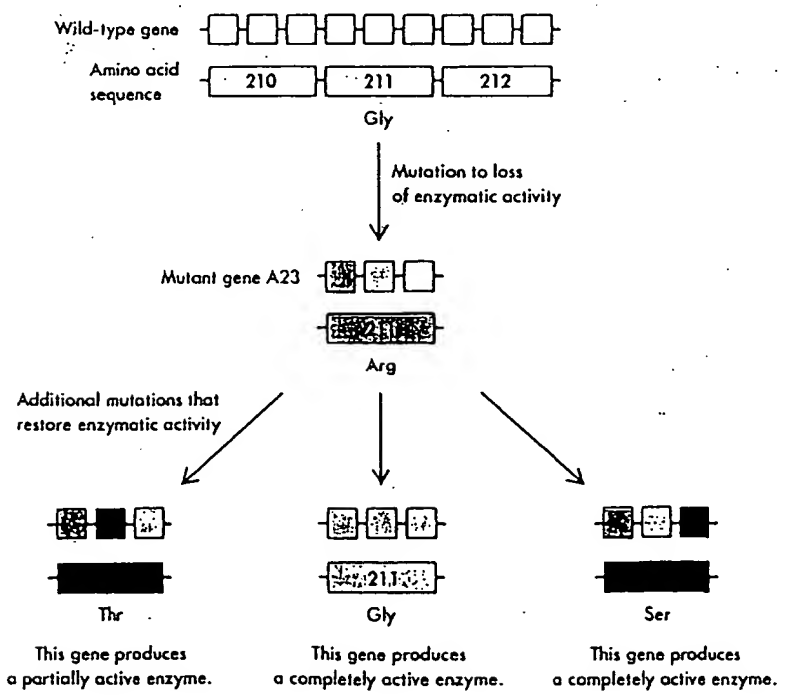


Figure 8-16
Evidence that many amino acid replacements do not result in loss of enzymatic activity.

A Second Amino Acid Replacement May Cancel Out the Effect of the First[10]

The conclusion that minor changes to amino acid sequence do not significantly alter enzyme activity is extended by the finding that some mutations that convert inactive mutant enzymes to active forms may work by causing a second amino acid replacement in the mutant enzyme. Consider mutant A46, which produces inactive tryptophan synthetase because of the substitution of glutamic acid for glycine at protein 212. In this case, distant second-site mutations that result in the active enzyme occasionally emerge. For example, the second-site mutation A446 is located one-tenth of a gene length away from the first mutation. The double mutant A46A446 produces active enzyme molecules containing two amino acid replacements: the original glycine-to-glutamic acid shift and a tyrosine-to-cysteine shift located 36 amino acids away (Figure 8-17).

The second shift can be studied independently of the first by obtaining recombinant cells with only the A446 mutation. Most interestingly the A446 change, when present alone, also results in an inactive enzyme. We thus see that a combination of two wrong amino acids can produce an enzyme with an active three-dimensional configuration. However, only occasionally do two wrong amino acids cancel out each other's faults. For example, double mutants containing A446 and A23, or A446 and A187, do not produce active enzyme. At this time, it does not seem wise to speculate on how the various amino acid residues are folded together in the three-dimensional configuration and why only some combinations are enzymatically active. This kind of analysis must await the establishment of the three-dimensional structure of tryptophan synthetase.

The Very Drastic Consequences of the Insertion or Deletion of Single Base Pairs[11,12]

Early on in the analysis of mutant proteins, it became clear that the vast majority of mutants being isolated did not yield the minimally altered proteins, bearing single amino acid replacements, that would arise through the change of one type of base pair into one of its three alternatives. Instead, most mutants represented changes that led to drastically altered gene products, often containing many fewer amino acids and with many of their amino acid sequences bearing no relationship to the wild-type polypeptide products. The nature of these mutants first became apparent through the proposal that such mutations usually represented either insertions or deletions of single nucleotide pairs. The drastic effect of these insertion or deletion events is a consequence of the fact that mRNA molecules are read in successive blocks of three nucleotides, called codons. AUG codons, which code for the methionine residues found at the amino terminal ends of newly synthesized polypeptide chains, are the signal for ribosomes to begin reading the mRNA molecule about to be translated into a protein. Since reading always begins at the appropriate AUG condon, the mRNA molecules are aligned on the ribosomes so that their messages are read in the correct reading frame.

If, however, a single base pair is inserted or deleted in a coding sequence, the triplets that designate amino acids become completely changed beginning at the site of insertion or deletion (Figure 8-18).

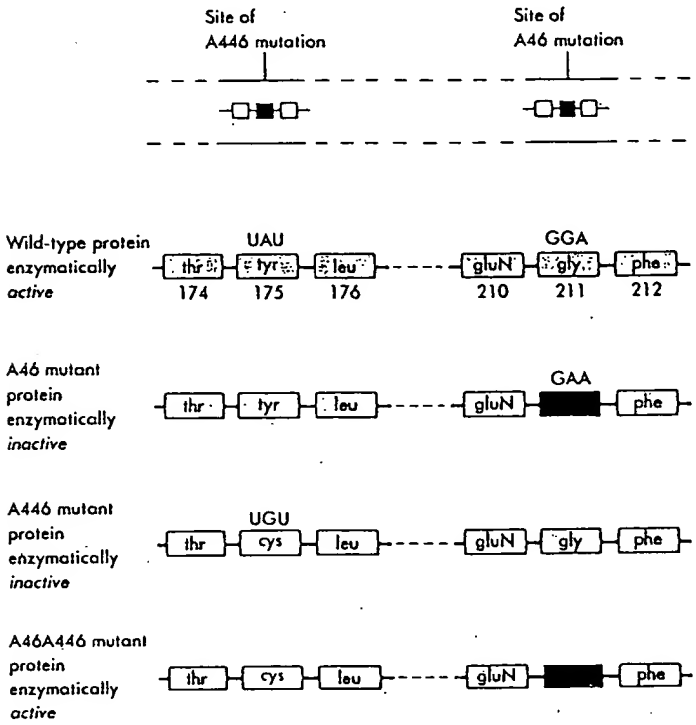


Figure 8-17
Reversal (suppression) of mutant phenotype by a second mutation at a second site in the same gene.

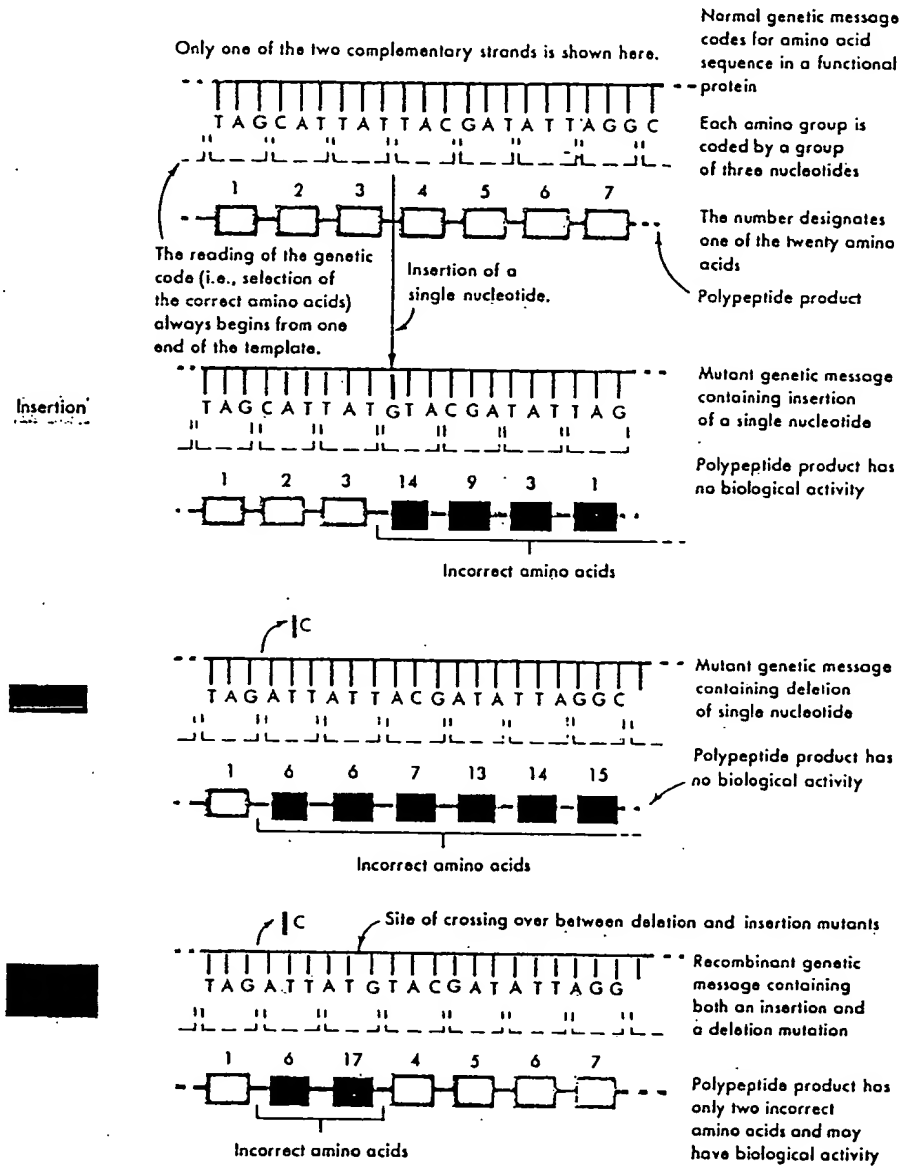
For example, if normally the gene sequence ATTAGACAC . . . is read as (ATT)(AGA)(CAC) . . . , then the insertion of a new nucleotide C in the fourth position of that sequence creates ATTCAGACAC, which is read as (ATT)(CAG)(ACA)(C . . .). These new triplets may code for entirely different amino acids. A similar consequence follows from a deletion. Moreover, the crossing of two deletion or two insertion mutants yields double mutants in which the reading frame is still misplaced.

Reversion of Insertion or Deletion Mutants

Active (or partially active) genes are regenerated by crossing over between an insertion and a nearby deletion. Such events restore the correct reading frame except in the short region between the mutations (see Figure 8-18). If the affected gene region is nonessential (e.g., the early section of the T4 *rIIB* gene), then the resulting protein product is fully functional. In other cases, the short segments of inappropriate amino acids are only mildly disadvantageous, and partial activity results. No activity, however, will usually be found if the inappropriate codons include any of the three that signify chain termination (UAA, UAG, or UGA). Their presence inevitably results in incomplete fragments of the wild-type polypeptide.

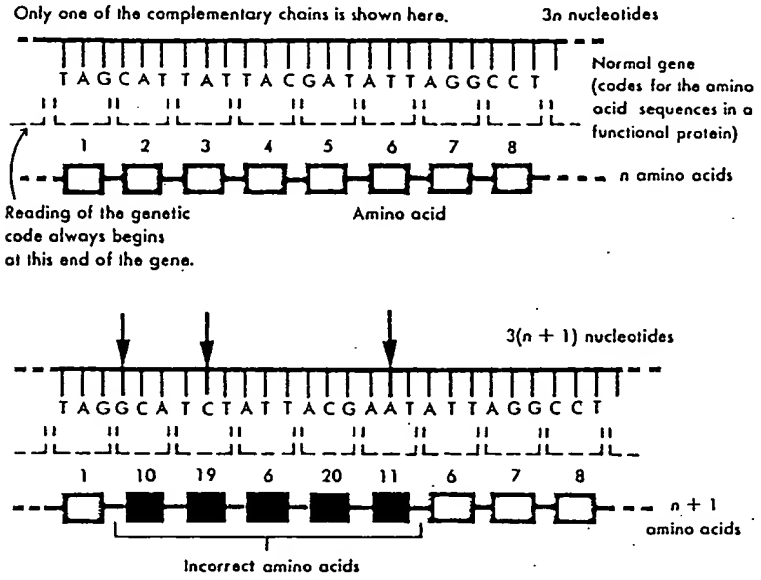
It is also sometimes possible to obtain functional genes by producing recombinants containing three closely spaced insertions or deletions (Figure 8-19). In contrast, recombinants containing four nearby insertions or deletions produce only nonfunctional polypeptides. These later experiments were performed in 1961, before the basic outlines of the genetic code were known. They in fact provided the first good evidence that the genetic code was likely to be read in groups of three as opposed to groups of two or four.

Figure 8-18
Mutations that add or remove a base shift the reading frame of the genetic message.



Cloned Genes Can Be Sequenced[13–17]

Virtually all the essential features of the genetic code were deduced by 1966 from the coding properties of either enzymatically or chemically synthesized mRNA molecules and from the accumulated knowledge of genetic fine structure that we have just detailed. No real genes were directly analyzed, however, since at that time there were no procedures either to sequence DNA or to isolate desired genes. But with the arrival of recombinant DNA and of powerful methods for DNA sequencing, the nature of genetic research has dramatically changed. No longer are genetic crosses the prime vehicle for probing genes. The quickest and most direct way to proceed is now the cloning and sequencing of relevant genetic material. As indicated in the previous chapter, it is now a relatively straightforward matter to isolate any *E. coli* gene that codes for a function that can be selected for by one of the many enrichment procedures.



Polypeptide chain contains five incorrect amino acids; its chain length is increased by one amino acid. It may have some biological activity depending upon how the five wrong amino acids influence its 3-D structure.

Figure 8-19
When three nucleotides are added close together, the genetic message is scrambled only over a short region. The same type of result is achieved by the deletion of three nearby nucleotides.

Already, a large number of *E. coli* genes have been completely or partially sequenced. In all cases, the codons found to specify given amino acids are those predicted by the genetic code (Figure 8-20). This agreement between prediction and result, though inherently very satisfying, surprised no one, since the experimental evidence used to deduce the genetic code was effectively unassailable (see Chapter 15). Also as predicted, the coding segments of virtually all mRNAs start with the AUG codon and always conclude with a chain-terminating codon (UAA, UAG, or UGA).

Untranslated Sequences at the Beginnings and Ends of mRNA Molecules[18–23]

When mRNA was first discovered, it seemed simplest to assume that the translation events would begin at one end of the molecule and then move along in steps of three nucleotides until the other end was reached. This was a very naive view, adopted before the discoveries that methionine initiates all polypeptide chains and that specific codons specify chain termination. Now we realize that untranslated sequences exist at both the 5′ end of the mRNA, near which translation begins, and at the 3′ end, near which translation stops (Figure 8-21). Hence, there must be internal signals in mRNA that mark the starting and stopping sites for translation. With the exception of a small purine-rich block of nucleotides that functions to position ribosomes at the correct AUG start codon, the untranslated regions probably play no role in translation and are of variable lengths, ranging from 20 to more than 100 nucleotides, depending on the particular mRNA species.

These seemingly unnecessary extra sequences only make sense

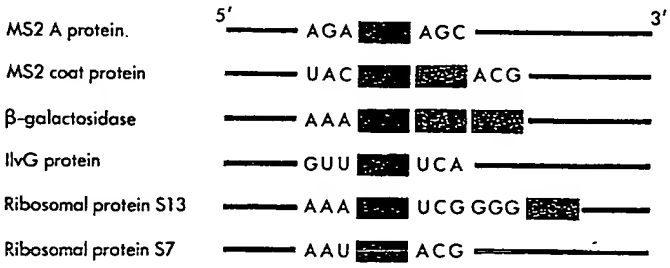


Figure 15-8
Nucleotide sequences at the 3′ ends of coding regions of mRNAs translated in *E. coli*. The stop codons are indicated by color boxes. Tandem stop codons do occur, but are rare. Likewise, when many mRNAs are compared, the average distance from the terminator to the next in-frame stop codon is about as expected on a random basis. [After J. Kohli and H. Grosjean, *Mol. Gen. Genetics* 182 (1981):430.]

entially used to end polypeptide coding regions. Now that the nucleotide sequences at the ends of a number of *E. coli* genes have been elucidated, it is clear that UAA is the preferred, but by no means exclusive, signal. In most cases, only a single stop codon appears. But some genes end with two or even three successive stop signals (Figure 15-8). The presence of more than one stop codon may be a precaution against the rare case in which the first codon fails; but why this device is used only occasionally is unclear. For example, in the RNA phages, homologous coat protein genes terminate with either one stop codon (in phage Qβ) or two stop codons (in phages R17 and MS2).

Nonsense Versus Missense Mutations[17, 18]

An alteration that changes a codon specific for one amino acid to a codon specific for another amino acid is called a **missense mutation**. The change to a chain-termination codon is known as a **nonsense mutation**. Given the existence of only three chain-termination codons, most mutations involving single-base replacements (**point mutations**) are likely to result in missense rather than nonsense. Since a new protein arising by missense mutation contains only a single amino acid replacement, it frequently possesses some of the biological activity of the original protein. Often, missense proteins fail to function only at higher than normal temperatures and are therefore known as **temperature-sensitive mutations** (Chapter 7). Many of the abnormal hemoglobins (see Figure 3-12) are the result of missense mutations. Amino acid replacement data obtained from these changed hemoglobin molecules strongly support the idea that these mutations result from the substitution of single nucleotides.

Nonsense Mutations Produce Incomplete Polypeptide Chains[17, 18]

When a nonsense mutation occurs in the middle of a genetic message, an incomplete polypeptide is released from the ribosome owing to premature chain termination. The size of the incomplete polypeptide chain depends on the location of the nonsense mutation. Mutations occurring near the beginning of a gene result in very short fragments, while mutations near the end produce fragments of almost normal length. Most incomplete chains have no biological activity, making most nonsense mutations in vital genes easily detectable. In contrast, the majority of missense mutations have some biological activity and can be easily overlooked. Thus, after treating *E. coli* with

Suppressor Mutations Can Reside in the Same or a Different Gene[17–21]

Often, the effects of harmful mutations can be reversed by a second genetic change. Some of these subsequent mutations are very easy to understand, being simple **reverse** (back) **mutations**, which change an altered nucleotide sequence back to its original arrangement. Much more difficult to understand are the mutations occurring at different locations on the chromosome that suppress the change due to a mutation at site A by producing an additional genetic change at site B. Such **suppressor mutations** fall into two main categories: those occurring within the same gene as the original mutation but at a different site in this gene (**intragenic suppression**) and those occurring in another gene (**intergenic suppression**). Genes that cause suppression of mutations in other genes are called **suppressor genes**.

Now we realize that both types of suppression work by causing the production of good (or partially good) copies of the protein made inactive by the original harmful mutation. For example, if the first mutation caused the production of inactive copies of one of the enzymes involved in making arginine, then the suppressor mutation allows arginine to be made by restoring the synthesis of some good copies of this same enzyme. However, the mechanisms by which intergenic and intragenic suppressor mutations cause the resumption of the synthesis of good proteins are completely different.

Those mutations that can be reversed through additional changes in the same gene often involve insertions or deletions of single nucleotides. These shift the reading frame (Chapter 8) so that all the codons following the insertion (or deletion) are completely changed, thereby generating new amino acid sequences. Often, the shifted reading frame will contain nonsense codons, and as a result, prematurely terminated polypeptides will be produced by the mutant cell. Intragenic suppression may occur when a second mutation deletes (or inserts) a new nucleotide near the original change and thus restores the original codon arrangement beyond the second change (Figure 15-9). Even though there are still scrambled codons between

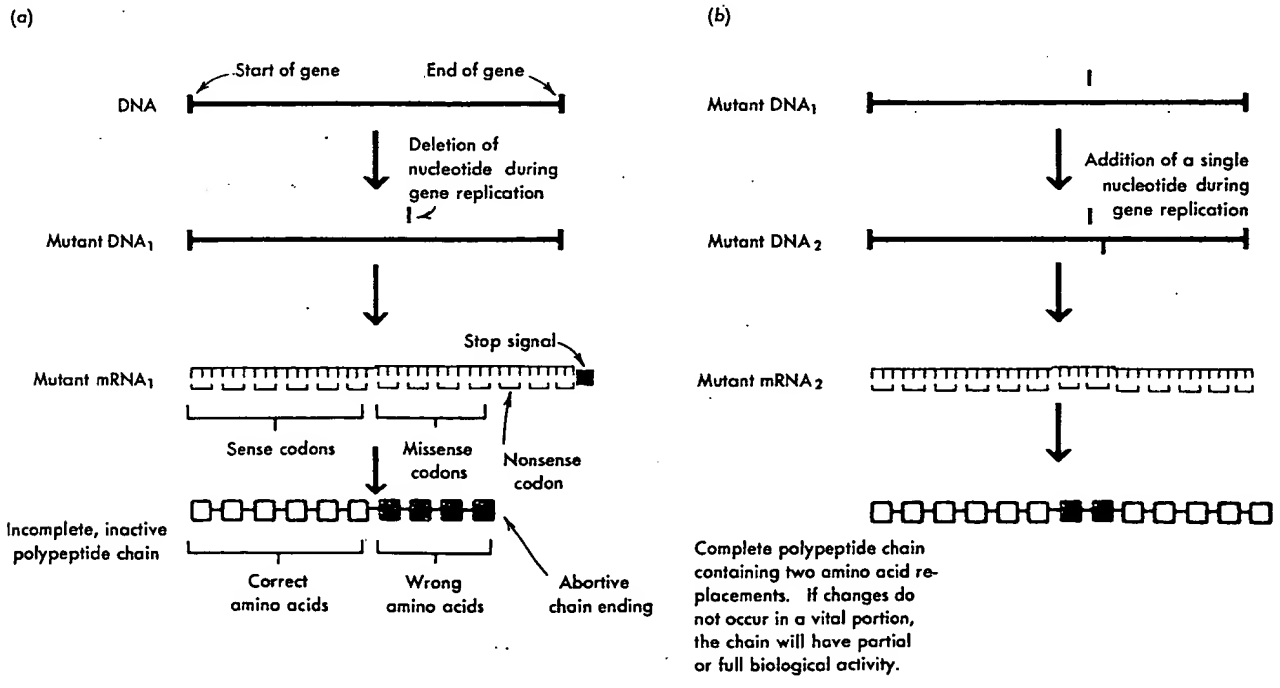


Figure 15-9
Intragenic suppression of a nucleotide deletion or insertion mutation. (a) The effects of a single-nucleotide deletion mutation upon the reading of the genetic message. (b) The mechanism by which a nucleotide addition mutation can suppress the havoc caused by the previous deletion mutation. In a similar way, the effect of a nucleotide addition could be overcome by a subsequent nucleotide deletion.

the two changes, there is a good probability, because of degeneracy, that the scrambled codons all code for some amino acid. If so, full-length, often functional proteins may be produced.

When the original mutation is a missense mutation, intragenic suppression can also result from a second missense mutation. In these cases, the original loss of enzymatic activity is due to an altered three-dimensional configuration resulting from the presence of a wrong amino acid. A second missense mutation in the same gene brings back biological activity if it somehow restores the original configuration around the functional part of the molecule. An example of this type of suppression in the tryptophan synthetase system was shown in Chapter 8 (see Figure 8-17).

Suppressor Genes Upset the Reading of the Genetic Code[17, 18, 20, 21]

Suppressor genes do not act by changing the nucleotide sequence of a mutant gene. Instead, they change the way the mRNA template is read. There are a number of different suppressor genes in *E. coli*. Since each causes the misreading of a specific nonsense or missense codon, suppressor genes can reverse the effects of only a small fraction of the point mutations that might arise within a given gene. For example, if we collect a large number of mutations blocking the synthesis of the enzyme β-galactosidase (Chapter 16), only several percent of these mutations will be suppressed by suppressor gene *a*. These few mutations will have nucleotide replacements in codons whose reading is specifically altered by gene *a*. Similarly, a completely different small fraction of β-galactosidase mutations can be suppressed by suppressor gene *b*. Thus, we see that specific codons are misread by specific suppressor genes.

On the other hand, since each suppressor gene causes the misreading of a specific codon, it is easy to understand how a given suppressor gene can suppress mutations in a number of different protein-coding genes. For example, the ability to synthesize both arginine and tryptophan in certain double mutants unable to make either amino acid can be restored by the presence of a single suppressor gene. We merely need to postulate that both these growth requirements are caused by the same specific codon change to missense or nonsense.

Nonsense Suppression Involves Mutant tRNAs[21–27]

There are suppressor genes for each of the three chain-terminating codons. They act by reading a stop signal as if it were a signal for a specific amino acid. There are, for example, three well-characterized genes that suppress the UAG codon. One suppressor gene inserts serine, another glutamine, and a third tyrosine at the nonsense position. In each of the three UAG suppressor strains, the anticodon of a tRNA species specific for one of these amino acids has been altered. For example, the tyrosine suppressor arises by a mutation within a $tRNA^{Tyr}$ gene that changes the anticodon from 3′-AUG-5′ to 3′-AUC-5′, thereby enabling it to recognize UAG codons (Figure 15-10). The serine and glutamine suppressor tRNAs also arise by single-base